

Brain networks of affective mentalizing revealed by the tear effect: The integrative role of the medial prefrontal cortex and precuneus

Haruka K. Takahashi^{a,b}, Ryo Kitada^{a,b,*}, Akihiro T. Sasaki^{c,d}, Hiroaki Kawamichi^{a,e,f}, Shuntaro Okazaki^a, Takanori Kochiyama^g, Norihiro Sadato^{a,b,*}

^a Division of Cerebral Integration, National Institute for Physiological Sciences, Okazaki 444-8585, Japan

^b Department of Physiological Sciences, SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193, Japan

^c Pathophysiological and Health Science Team, RIKEN Center for Life Science Technologies, Kobe 650-0047, Japan

^d Department of Physiology, Osaka City University Graduate School of Medicine, Osaka 545-8585, Japan

^e Graduate School of Human Health Sciences, Tokyo Metropolitan University, Tokyo 116-8551, Japan

^f School of Medicine, Faculty of Medicine, Gunma University, Maebashi 371-8511, Japan

^g ATR Brain Activity Imaging Center, Seika-cho 619-0288, Japan

ARTICLE INFO

Article history:

Received 5 February 2015

Received in revised form 29 April 2015

Accepted 6 July 2015

Available online 18 July 2015

Keywords:

Empathy

Expression

Face

fMRI

Mentalizing

Tear

ABSTRACT

Affective mentalizing involves the integration of various social signals in order to infer the affective states of others. Previous neuroimaging studies have shown that the medial prefrontal cortex, the precuneus/posterior cingulate cortex, and the temporo-parietal junction constitute the core affective mentalizing network. However, the relative contributions of these regions to affective mentalizing remain unclear. We used functional magnetic resonance imaging to investigate which of these nodes are involved in the integration of two social signals: emotional tears and facial expressions. We assumed that this integration would produce a supra-additive effect, indicated by greater activity than the sum of the effects of the individual social signals. Female subjects rated the sadness of faces with either tears or tear-like circles, and either sad or neutral expressions. We observed the supra-additive effect in the medial prefrontal cortex and precuneus/posterior cingulate cortex but not in the temporo-parietal junction. These results indicate that the medial prefrontal cortex and precuneus/posterior cingulate cortex play an important role in integrating tears and facial expressions during affective mentalizing.

© 2015 The Authors. Published by Elsevier Ireland Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Social cognition involves psychological processes that allow humans to make inferences about other people. Among such processes, affective mentalizing (or cognitive empathy) is defined as the process of inferring another's affective state (i.e., "I understand how you feel") (Perry and Shamay-Tsoory, 2013). Previous neuroimaging and lesion studies have identified a number of brain regions involved in affective mentalizing (Frith and Frith, 2003; Samson et al., 2004; Saxe and Powell, 2006; Mitchell, 2008; Shamay-Tsoory et al., 2009; Atique et al., 2011; Corradi-Dell'Acqua et al., 2014). Among these regions, the medial prefrontal cortex (mPFC), the precuneus/posterior cingulate cortex (PCC), and the temporo-parietal junction (TPJ) are considered the core

mentalizing network, because they have often been observed during affective (Atique et al., 2011; Corradi-Dell'Acqua et al., 2014) and non-affective (cognitive) (Goel et al., 1995; Van Overwalle, 2009; Van Overwalle and Baetens, 2009) contexts. However, the relative contributions of these nodes to affective mentalizing are not well understood.

One possible way to clarify the relative contributions of these nodes is to examine the brain regions that are involved in integrating multiple social signals to infer another's affective state. Here, we define the integration as a process in which social signals are combined to infer the most likely affective state. In the field of multisensory research, if two different types of social signal are integrated in a region, it is expected not only to be activated by each separately (convergence), but also to show interaction effects between them (Calvert et al., 2000; Rajj et al., 2000; Stevenson et al., 2009). Previous neuroimaging studies have consistently found that the mPFC contains information about another's emotional state, regardless of whether the social signals involved facial, body, or vocal expressions (Peelen et al., 2010), or whether they involved facial expressions or situational information in the absence of

* Corresponding authors at: Division of Cerebral Integration, National Institute for Physiological Sciences, Okazaki 444-8585, Japan. Tel.: +81 564 55 7841; fax: +81 564 55 7843.

E-mail addresses: kitada@nips.ac.jp (R. Kitada), sadato@nips.ac.jp (N. Sadato).

observable expressions (Skerry and Saxe, 2014). These findings indicate that the mPFC plays a key role in representing others' affective states at the abstract level by receiving information about distinct social signals. However, each type of social signal was presented separately in previous studies, so it has remained unclear whether the mPFC is involved in the integration process.

Even less is known about the role of the PCC and TPJ in the integration of social signals. In particular, the function of the PCC in mentalizing is not well understood; hence, little attention has been paid to its role in integrating social signals. Furthermore, the function of the TPJ in mentalizing has been controversial (Decety and Lamm, 2007; Mitchell, 2008; Scholz et al., 2009; Cabeza et al., 2012), and its role in the integration of social signals has not been clarified. For instance, Peelen et al. (2010) showed that a region adjacent to the TPJ also contains information about another's emotional state across different types of emotional expression (face, body, and voice). However, a subsequent neuroimaging study by Skerry and Saxe (2014) utilizing facial expressions and situational information (e.g., social exclusion) showed that the TPJ did not represent another's emotional state at an abstract level; more specifically, their work showed that a classifier, which was trained to discriminate the valence of one social signal within the TPJ, did not successfully classify that valence for the other social signals. Understanding another person's affective state based on situational information is critically different from interpreting emotional expressions (produced by the face, body, and voice), in that situational information can be interpreted in multiple ways and presents an ill-posed inverse problem (e.g., a person might feel happy or sad when he or she is separated from others). Thus, understanding situational information in a socially-appropriate manner requires knowledge of the social event (e.g., that separation from others should be considered a sad event; Barbey et al., 2009; Krueger et al., 2009). Thus, we anticipate that the TPJ plays a minor role in the integration of social signals that involve social event knowledge.

Given this background, we focused our investigation on the integration of two social signals from facial stimuli: tears and facial expressions. To the best of our knowledge, neither the neural mechanisms underlying the processing of tears as social stimuli nor the neural bases of the integration of tears and facial expressions has been identified. Emotional tears appear to be unique to humans and are of considerable interest in the field of evolutionary psychology (e.g., Murube et al., 1999; Provine et al., 2009; Balsters et al., 2013). Like social situations, understanding another's affective state from tears is an ill-posed inverse problem, because tears can be shed in response to many different emotions (e.g., anger, happiness, and sadness; Murube et al., 1999). Therefore, tears are similar to social situations in that they require social event knowledge in order to achieve the most appropriate interpretation. In the absence of contextual information, humans tend to interpret tears as a symbol of sadness (i.e., the tears effect; Provine et al., 2009), possibly because such an interpretation is the most socially appropriate. As the mPFC can represent others' emotions at an abstract level across different social signals (Peelen et al., 2010; Skerry and Saxe, 2014), we predicted that it would be involved in integrating tears and facial expressions.

The present study used functional magnetic-resonance imaging (fMRI) to test the hypothesis that the mPFC, but not the TPJ, integrates tears and facial expressions during the evaluation of others' sadness. We also explored the role of the PCC in this integration process without a specific hypothesis. We manipulated two factors: tears (tears, tear-like control objects, and no object) and facial expressions (sad and neutral). We initially tested our assumption that the core mentalizing network is activated by the presence of tears, and then examined whether this network shows interaction effects between tears and facial expressions. We predicted that

the mPFC would show a supra-additive effect, providing evidence of the integration of information on tears and facial expressions (Meredith and Stein, 1983; Calvert et al., 2000; Raij et al., 2000; Stevenson et al., 2009). In other words, these regions should show stronger activation in response to a sad facial expression with tears than the sum of the activity in response to individual presentations of a sad facial expression without tears and a neutral facial expression with tears. By contrast, we predicted that the TPJ would not show the same effect.

2. Materials and methods

2.1. Subjects

Sixty-one healthy subjects aged 18–44 years (mean age = 22.1 years; standard deviation [SD] = 4.7 years) participated in the study. We recruited only female participants because they tend to react to crying people with more sympathy and support than males (Cretser et al., 1982). All subjects were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). None of the volunteers had a history of symptoms requiring neurological, psychological, or other medical care. All subjects gave written informed consent. The study was approved by the ethical committee of the National Institute for Physiological Sciences of Japan. Thirty-eight subjects participated in the main fMRI experiment. The rest of the subjects ($n = 23$) participated in a separate experiment to define the regions of interest (ROIs). None of the subjects participated in both experiments.

2.2. Data acquisition

fMRI images were acquired using a 3T scanner (Verio; Siemens Erlangen, Germany) with a 32-element phased-array head coil. Tight but comfortable foam padding was placed around each subject's head to minimize movement. T2*-weighted gradient-echo echo-planar imaging (EPI) was used to obtain the functional images. The sequence parameters were as follows: repetition time (TR), 3000 ms; echo time (TE), 30 ms; flip angle, 83°; 39 slices of 3.0 mm thickness with a 17% slice gap, which covered the entire cerebral and cerebellar cortices; field of view, 192 mm; and in-plane resolution, 3.0 mm × 3.0 mm. Oblique scanning was used to exclude the eyeballs from the images. For anatomical imaging, a T1-weighted three-dimensional (3D) magnetization-prepared rapid-acquisition gradient echo (MP-RAGE) sequence was obtained (TR = 1800 ms; TE = 2.97 ms; flip angle = 9°; field of view = 250 mm; and voxel dimensions = 0.9 mm × 0.9 mm × 1.0 mm).

2.3. Stimuli





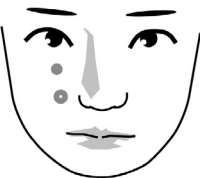

We used six types of face: those portraying sad facial expressions with tears, with tear-like circles, and without tears; and those portraying neutral facial expressions with tears, with tear-like circles, and without tears (Fig. 1A). Stimuli were produced as described below.

2.3.1. Stimuli production

We followed the same procedure as Provine et al. (2009) to produce the stimuli. We initially obtained 90 images of faces with tears (Tears images) from the online image archives Flickr (www.flickr.com) and Google (www.google.co.jp). We limited our search to images of female adults in order to eliminate gender differences between the subjects and stimuli. In addition to facial images, we also collected 45 landscape images from Flickr to use as controls.

We removed the tears from the 90 face images using photo-editing software (Adobe Photoshop, Adobe Systems Inc., San Jose, CA) and defined them as NoTears images. We then produced new

A. Task design

		Tears		
		With Tear (Tears)	With Circle (Circles)	Without Tear (NoTears)
Expressions	Sadness	 Sad + Tears (ST)	 Sad + Circles (SC)	 Sad (S)
	Neutral	 Neutral + Tears (NT)	 Neutral + Circles (NC)	 Neutral (N)

B. Task schedule (block design)

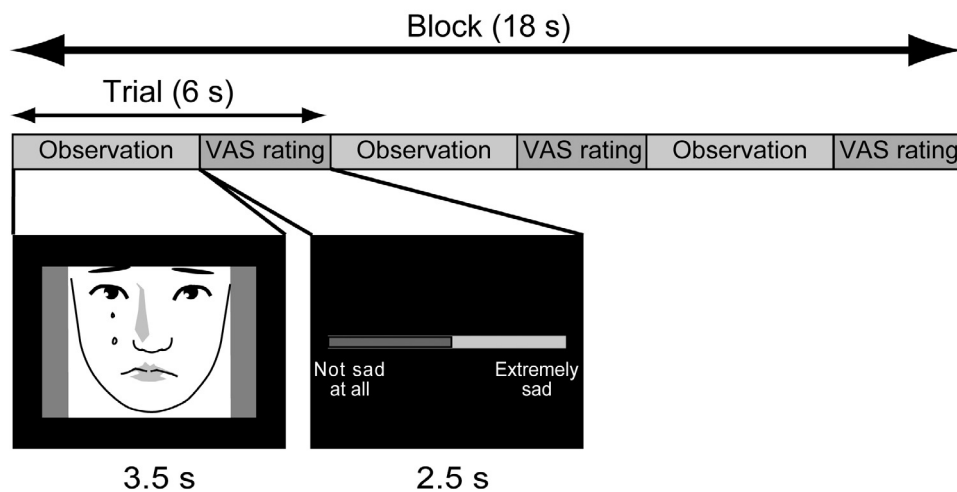


Fig. 1. Experimental design. (A) Task designs. We adopted a within-subjects factorial design with two factors: Tears and Facial expression. We extended the task design of [Provine et al. \(2009\)](#) by including two levels of facial expressions (Sad and Neutral) and three levels of Tears (Tears, Circles, and NoTears). Note that due to copyright issues these are schematics rather than the actual stimuli. (B) Task schedule. We used conventional block designs in which three trials of the same condition (with different pictures) were repeated three times in a block (18 s). The order of conditions was pseudo-randomized. In each trial (6 s), the subject viewed the face picture for 3.5 s and evaluated the extent of sadness of the person for 2.5 s.

images by adding gray circles to the NoTears images, resulting in the Circles images. The location and the number of gray circles in the Circles images were matched with the location and number of tears in the original (Tears) images. We did not include natural objects on stimuli (e.g., scars or saliva) in order to match the locations of stimuli and to avoid any possible interpretation of affective states (e.g., the observed person is hungry because of saliva on the face). Collectively, we created three sets of facial images (comprising 270 images in total): 90 Tears images, 90 NoTears images, and 90 Circles images. The mean differences in size and perceived brightness of the images for each condition were minimized using photo-editing software (Adobe Photoshop, Adobe Systems Inc., San Jose, CA).

We further categorized each set of images into two subsets (sad and neutral) based on their facial expressions. Eight females, who did not participate in the fMRI experiment, rated the intensity of sadness in the images on a visual analog scale (VAS) ranging

from 0 (“not sad at all”) to 100 (“extremely sad”). Initially, we used the VAS scores from the same eight subjects to classify the 90 NoTears images into 45 images of “sad” expressions and 45 images of “neutral” expressions. Then, the Tears and Circles images were categorized into “sad” and “neutral” images; the images for each facial expression were identical except for the presence of tears or circles. In total, we produced seven types of image: sad facial expressions with tears (Sad Expression + Tears [ST] images), sad facial expressions without tears (Sad Expression [S] images), sad facial expressions with circles (Sad Expression + Circles [SC] images), neutral facial expressions with tears (Neutral + Tears [NT] images), neutral faces without tears (Neutral [N] images), neutral faces with circles (Neutral + Circles [NC] images), and landscape images (Baseline [B] images) ([Fig. 1A](#)). [Fig. 1](#) shows schematic drawings of representative stimuli due to copyright issues. Each image was used only once in each experiment.

2.3.2. Stimulus presentation

Stimuli were back-projected via a liquid crystal display (LCD) projector (CP-SX12000; Hitachi, Ltd., Tokyo) onto a translucent screen located at the rear of the scanner. The horizontal and vertical viewing angles of stimuli were 5.3° and 7.4° , respectively. The subjects viewed stimuli via a mirror placed above the head coil. We used Presentation software to display visual stimuli and record the subject's response (Neurobehavioral Systems, Inc., San Francisco, CA).

2.4. Task design and procedure

A two-factor within-subjects factorial design was used, with two levels of Facial Expressions (Sad and Neutral) and three levels of Tears (Tears, NoTears and Circles) (Fig. 1A). In addition to these six conditions, we included a baseline (B) condition, during which subjects observed landscape images.

We employed a conventional block design (Fig. 1B) with five runs. Each run consisted of the first 12 scan volumes, followed by 23 blocks that lasted for 18 s (6 volumes per block), and the final 6 volumes (12 volumes + [23 blocks \times 6 volumes] + 6 volumes = 156 volumes per run). Each block included one of the seven task conditions, and each condition was repeated three times (21 blocks). Each block included three trials of the same condition, and each trial lasted 6 s (3 blocks \times 3 trials per block \times 5 runs = 45 trials for each condition in total). The order of the conditions was pseudo-randomized in each repetition. In addition, we included two blocks of the rest condition: one between the last block of the first repetition of a condition and the first block of the second repetition; and the other between the last block of the second repetition and the first block of the third repetition (2 rest blocks + 21 task blocks = 23 blocks in total).

In each trial, an image was presented for 3.5 s, followed by the presentation of a visual analog scale (VAS) for 2.5 s. The subjects used the VAS to evaluate the extent of sadness of the presented facial stimuli. The subjects manipulated a two-button response box (HHSC-2 \times 2, Current Designs, Inc., Philadelphia, PA) with their right hand to specify the location of the vertical line on a VAS (the index finger moved the line to the left, and the middle finger moved the line to the right). The VAS scale consisted of a white-colored horizontal bar with each end indicating the minimum (i.e., not sad at all) and the maximum (i.e., extremely sad) of the intensity of sadness expressed by the image. The end of the minimum and maximum was counterbalanced across the subjects. In the baseline (B) condition, the subject was asked to move the VAS to any position they wished. All of the subjects performed several practice trials in order to familiarize them with the task and to ensure they were able to utilize the VAS easily.

2.5. Imaging data processing

The first six volumes of each fMRI run were discarded for stabilization of the magnetization, and the remaining 150 volumes per run (a total of 750 volumes per participant) were used for analysis. Image processing and statistical analyses were performed using the Statistical Parametric Mapping (SPM8) package (Friston et al., 2007). The images were realigned to correct for head motion, then corrected for differences in slice timing within each volume. After the T1-weighted anatomical images were segmented into different tissue classes, each subject's T1-weighted anatomical image was co-registered with the mean image of all of the EPI images for each subject. Each co-registered T1-weighted anatomical image was normalized to the Montreal Neurological Institute T1 image template (ICBM 152) (Evans et al., 1994; Friston et al., 1995). The parameters from this normalization process were then applied to each functional image. The spatially normalized EPI images were

filtered using a Gaussian kernel of 8 mm full width at half maximum (FWHM) in the x, y, and z axes (final smoothness: $x = 11.8$, $y = 11.9$, and $z = 11.8$ mm). The parameters from this normalization process were then applied to the functional images, which were resampled to a final resolution of $2 \times 2 \times 2$ mm³.

2.6. Statistical analyses

2.6.1. Individual analyses

A design matrix comprising the five runs was prepared for each subject. We fitted a general linear model (GLM) to the fMRI data for each subject (Friston et al., 1994a; Worsley and Friston, 1995). Neural activity during each condition was modeled with box-car functions convolved with the canonical hemodynamic-response function. Each run included seven task-related regressors, one for each condition. The time series for each voxel was high-pass filtered at 1/128 Hz. Assuming a first-order autoregressive model, the serial autocorrelation was estimated from the pooled active voxels with the restricted maximum likelihood (ReML) procedure, and was used to whiten the data (Friston et al., 2002). Motion-related artifacts were minimized by incorporating six parameters (three displacements and three rotations) from the rigid-body realignment stage into each model. The parameter estimates for each condition in each individual were compared using linear contrasts. After confirming face-related activation (e.g., activation in the fusiform gyrus) by comparing face conditions with baseline (B), we evaluated the following contrasts (Table 1): first, Tears minus NoTears, [(ST + NT) – (S + N)]; second, Tears minus Circles, [(ST + NT) – (SC + NC)]; third, Sad minus Neutral expressions, [(ST + S + SC) – (NT + N + NC)]; and fourth, the interaction effects, [(ST – SC) – (NT – NC)] and [(NT – NC) – (ST – SC)]. The supra-additive effect was tested by the contrast of [(ST – SC) – (NT – NC)] (i.e., (ST – NC) > (SC – NC) + (NT – NC)).

2.6.2. Random-effects group analysis

In the first-level analysis, we obtained images that represent the normalized task-related increment of the MR signal of each subject for each predefined contrast (i.e., contrast images). These contrast images were used for the group analysis. For each predefined contrast (Table 1), a one-sample *t*-test was performed for every voxel in the brain to obtain population inferences (Holmes and Friston, 1998). The resulting set of voxel values for each comparison constituted an SPM of the *t* statistic SPM{*t*}. The height threshold for the SPM{*t*} was set at $t(37) > 2.72$, equivalent to $p < 0.005$ uncorrected. The statistical threshold for the spatial extent test on the clusters was set at $p < 0.05$ and corrected for multiple comparisons (family-wise error [FWE]) over the whole brain (Friston et al., 1994b, 1996).

Brain regions were anatomically defined and labeled according to probabilistic atlases (Amunts et al., 2005; Eickhoff et al., 2005; Shattuck et al., 2008) and a previous meta-analysis study (Van Overwalle, 2009). In order to avoid the ambiguity of the anatomical location of the TPJ, we defined it as the angular gyrus (Saxe and Powell, 2006; Scholz et al., 2009; Cabeza et al., 2012). Consistent with a meta-analysis (Van Overwalle, 2009), we defined the mPFC as the medial wall of the prefrontal cortex: that is, regions in which the *x* coordinates ranged from -20 to 20 and the *y* coordinates were above $y > 20$ in MNI space (Van Overwalle, 2009). We further subdivided the mPFC into three regions (Van Overwalle, 2009): the dorsomedial prefrontal cortex (dmPFC), which lies above the *z* coordinate of 20 mm; the ventromedial prefrontal cortex (vmPFC), which lies between $z = -15$ and $z = 20$ mm; and the medial orbitofrontal cortex (mOFC), which lies below $z = -15$ mm.

2.6.3. ROI analysis

After the whole-brain analysis was completed, we conducted ROI analysis. The assumption was that the supra-additive effect of

Table 1
Predefined contrasts.

Regressors	Sad			Neutral			Baseline (B)
	ST	SC	S	NT	NC	N	
Tears – NoTears	1	0	–1	1	0	–1	0
Tears – Circles	1	–1	0	1	–1	0	0
Sad – Neutral expressions	1	1	1	–1	–1	–1	0
Interaction effects	1	–1	0	–1	1	0	0
	–1	1	0	1	–1	0	0

tears and facial expressions should be observed within the regions that were more active during the Tears condition than the NoTears condition, which is a more liberal control than the Circles conditions (SC and NC). Thus, we defined ROIs based on the brain regions that were activated by the Tears condition relative to the NoTears condition.

The use of the same dataset for the definition of ROI and analysis of response patterns in the ROI can lead to invalid statistical inferences (i.e., the double-dipping problem; Kriegeskorte et al., 2009). To avoid this, we conducted a separate experiment to compare the Tears condition with the NoTears condition, as described below.

The design and analyses of this experiment were identical to the main experiment except that the Circles conditions (SC and NC) were removed (3 blocks for each condition \times 5 conditions + 2 rest blocks = 17 blocks). We collected 9 volumes before the first block and 7 volumes after the last block ($9 + [17 \text{ blocks} \times 6] + 7 = 118$ volumes per run). The threshold for the SPM{t} was set at $t(22) > 2.82$ (equivalent to $p < 0.005$ uncorrected, which was the same height threshold as that in the main experiment). The statistical threshold for the spatial extent test on the clusters was set at $p < 0.05$ and corrected for multiple comparisons (FWE) over the whole brain (Friston et al., 1994b, 1996).

We evaluated only the Tears minus NoTears contrast [(ST + NT) – (S + N)]. We chose the peak coordinates in each cluster in the mPFC, PCC, and TPJ. A cluster of activation could include anatomical regions beyond our hypothesis (e.g., the fusiform gyrus in the TPJ ROI). In order to limit the ROIs to each hypothesized region, we calculated the overlap between the cluster activated by the Tears effect and a 12-mm-radius sphere with the peak coordinates of the same cluster. This radius was identical to the effective resolution (final smoothness) of the statistical parametric maps. This overlapping region in each cluster was used as the ROI.

We analyzed the data from the main experiment based on these ROIs. We used unsmoothed data in order to characterize activation patterns without blurring and to maximize the sensitivity. We averaged the contrast estimates in all voxels within each ROI and compared them between the four conditions of interest (ST, SC, NT, and NC).

3. Results

3.1. Behavioral results

The presence of tears on the face images increased the VAS ratings of sadness (Fig. 2). A two-way repeated-measures analysis of variance (ANOVA) (2 levels of Facial Expressions \times 3 levels of Tears) on the rating scores revealed significant main effects of Facial Expressions [$F(1, 37) = 635.8$, $p < 0.001$] and of Tears [$F(2, 74) = 142.5$, $p < 0.001$], and a significant interaction [$F(2, 74) = 94.2$, $p < 0.001$]. Post hoc pair-wise comparisons (with a Bonferroni correction) showed that there were greater VAS ratings in the Tears condition compared with the NoTears and Circles conditions for each facial expression (p values < 0.001). The effect of tears on the VAS ratings was greater for the neutral expressions

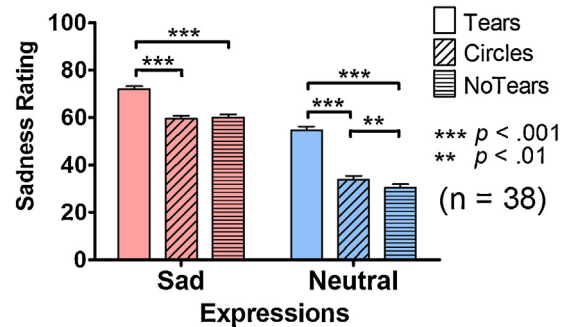


Fig. 2. Behavioral results. VAS ratings of sadness perceived by the subjects. Data are presented as the mean \pm standard error of the mean (SEM). Asterisks indicate significant differences revealed by post hoc pairwise comparisons (with the Bonferroni correction).

than for the sad expressions, regardless of whether the Tears condition was compared to the NoTears or Circles conditions (p values < 0.001). Finally, we found that the Circles condition showed greater VAS rating scores than the NoTears condition for neutral expressions ($p < 0.01$), but not for sad facial expressions ($p > 0.9$).

Taken together, these findings confirm that the presence of tears increased the sadness ratings (Provine et al., 2009).

3.2. fMRI results

3.2.1. Whole-brain analysis

3.2.1.1. The main effect of tears. We conducted the two contrasts to evaluate the brain regions by the presence of tears: that is, the contrast with a liberal control (NoTears) and a stringent control (Circles). The contrast of Tears minus NoTears [(ST + NT) – (S + N)] revealed regions of significant activation bilaterally in the vmPFC, mOFC, posterior cingulate gyrus, TPJ, superior parietal lobule, superior, middle and inferior occipital gyri, middle and inferior temporal gyri, fusiform gyrus, and caudate nucleus (Fig. 3 and Table 2). In addition, the same contrast revealed significant activation in the left hemisphere: specifically, in the dmPFC, precuneus, middle frontal gyrus, superior temporal gyrus, amygdala, hippocampus and brainstem.

The contrast of Tears minus Circles [(ST + NT) – (SC + NC)] revealed no significant activation.

3.2.1.2. The main effect of sad expression. The contrast of Sad minus Neutral expressions [(ST + SC + S) – (NT + NC + N)] revealed significant activation bilaterally in the mPFC (dmPFC, vmPFC, and mOFC), precuneus, and posterior cingulate gyrus (Fig. 4 and Table 3). In addition to these regions, the same contrast revealed bilateral activation in the postcentral gyrus, superior frontal gyri, precentral gyrus, superior parietal lobule, cuneus, caudate nucleus, putamen, superior and middle temporal gyrus, parahippocampal gyrus, insula, hippocampus, amygdala, and fusiform gyrus. We also found activation in the left cerebellum, left inferior temporal gyrus, right supramarginal gyrus, and right lingual gyrus.

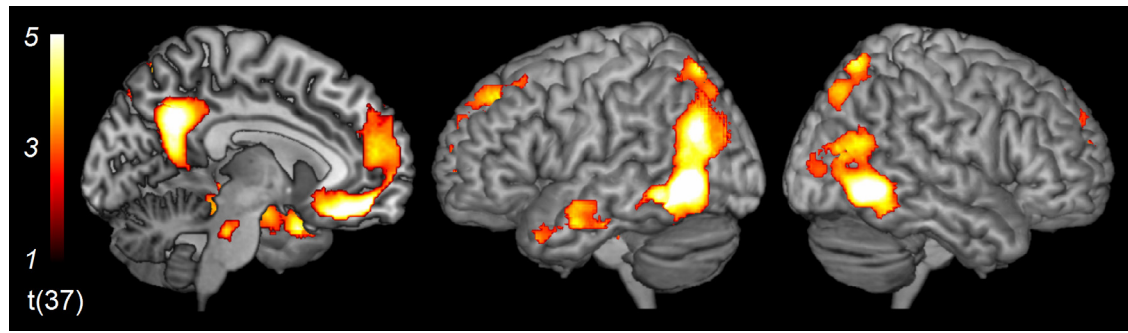


Fig. 3. Tears minus NoTears. The brain activation revealed by the Tear effect, via the contrast of Tears minus NoTears [(ST + NT) – (S + N)]. The activation was thresholded at $p < 0.05$, corrected for multiple comparisons over the whole brain, with the height threshold set at $t(37) > 2.72$ (corresponding to an uncorrected $p < 0.005$). The activation patterns were superimposed on surface-rendered high-resolution MRIs unrelated to the subjects of the present study.

Table 2

Tears minus NoTears.

Spatial extent test		MNI coordinates (mm)			t-Value(37)	Location	
Cluster size (mm ³)	p-Values	x	y	z		Side	Area
Tears minus NoTears [(ST + NT) – (S + N)] (Fig. 3) 37,656	<0.001	–26	–70	60	4.63	L	Superior parietal lobule
		–24	–80	46	3.11	L	Superior occipital gyrus
		–46	–68	24	5.82	L	TPJ
		–50	–64	–2	7.36	L	Middle temporal gyrus
		–48	–60	–4	7.55	L	Inferior temporal gyrus
		–50	–68	–4	6.12	L	Middle occipital gyrus
		–42	–60	–10	5.00	L	Fusiform gyrus
		–50	–70	–10	4.60	L	Inferior occipital gyrus
		–56	–6	–14	3.68	L	Superior temporal gyrus
		–26	–6	–18	3.90	L	Amygdala ^a
		–32	–28	–20	4.68	L	Hippocampus ^b
		–6	–28	–26	4.09	L	Brainstem
		–24	28	50	3.46	L	Middle frontal gyrus
		–16	38	44	4.19	L	dmPFC
		–8	20	–6	2.80	L	Caudate nucleus
17,872	<0.001	10	22	–6	3.21	R	Caudate nucleus
		–2	50	–12	6.03	L	vmPFC
		8	60	16	3.37	R	vmPFC
		0	26	–18	6.37		mOFC
		20	–66	64	4.56	R	Superior parietal lobule
17,528	<0.001	28	–78	46	3.51	R	Superior occipital gyrus
		58	–66	20	4.19	R	TPJ
		50	–76	18	3.93	R	Middle occipital gyrus
		52	–56	–4	5.95	R	Middle temporal gyrus
		52	–62	–6	6.20	R	Inferior temporal gyrus
		48	–64	–10	4.21	R	Inferior occipital gyrus
		46	–50	–16	4.08	R	Fusiform gyrus
		–4	–56	36	5.60	L	Precuneus
11,312	<0.001	–8	–54	26	5.33	L	Posterior cingulate gyrus
		4	–52	24	3.32	R	Posterior cingulate gyrus

The threshold size of activation was $p < 0.05$, corrected for multiple comparisons over the whole brain, when the height threshold was set at $t(37) > 2.72$. x, y, and z are stereotaxic coordinates (mm). R, right hemisphere; L, left hemisphere. dmPFC, dorsomedial prefrontal cortex; vmPFC, ventromedial prefrontal cortex; mOFC, medial orbitofrontal cortex; TPJ, temporoparietal junction. ^{a,b}Probability values on cytoarchitectonic maps (Amunts et al., 2005): ^a80% for the amygdala and 40% for the hippocampus; ^b50% for the hippocampus.

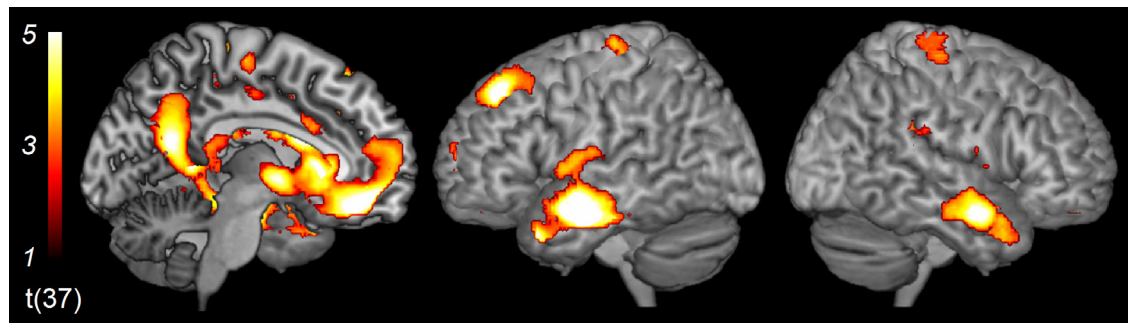


Fig. 4. Sad minus Neutral expressions. The brain activation revealed by sad expressions, via the contrast of sad minus neutral expressions [(ST + SC + S) – (NT + NC + N)]. The activation was thresholded at $p < 0.05$, corrected for multiple comparisons over the whole brain, with the height threshold set at $t(37) > 2.72$ (corresponding to an uncorrected $p < 0.005$).

Table 3
Sad minus Neutral expressions.

Spatial extent test		MNI coordinates (mm)			t-value(37)	Location	
Cluster size (mm ³)	p-Values	x	y	z		Side	Area
Sad minus Neutral expressions [(ST+SC+S) – (NT+NC+N)] (Fig. 4)							
167,592	< 0.001	–22	–22	68	3.92	L	Precentral gyrus
		40	–20	66	3.40	R	Precentral gyrus
		–22	–32	66	3.50	L	Postcentral gyrus
		40	–26	68	3.30	R	Postcentral gyrus
		–10	–18	60	3.83	L	Superior frontal gyrus
		14	–18	54	4.64	R	Superior frontal gyrus
		–20	36	42	6.20	L	dmPFC
		16	44	24	3.74	R	dmPFC
		–12	–56	32	3.87	L	Superior parietal lobule
		12	–56	24	3.68	R	Superior parietal lobule
		46	–38	24	3.65	R	Supramarginal gyrus
		–4	–52	20	4.60	L	Posterior cingulate gyrus
		6	–52	22	4.09	R	Posterior cingulate gyrus
		–4	–58	18	5.82	L	Precuneus
		4	–60	24	6.27	R	Precuneus
		–2	–66	16	3.07	L	Cuneus
		2	–68	24	3.97	R	Cuneus
		–12	20	6	5.41	L	Caudate nucleus
		4	20	4	4.80	R	Caudate nucleus
		–34	8	6	4.34	L	Insula
		34	8	8	4.91	R	Insula
		12	–44	4	3.05	R	Lingual gyrus
		–20	20	0	4.86	L	Putamen
		18	20	–4	4.01	R	Putamen
		–60	–2	0	3.29	L	Superior temporal gyrus
		48	–38	16	3.67	R	Superior temporal gyrus
		–4	–52	–4	3.36	L	Cerebellum
		–6	50	–10	5.65	L	vmPFC
		12	34	–10	5.03	R	vmPFC
		–30	–34	–10	4.94	L	Parahippocampal gyrus
		28	–20	–22	3.55	R	Parahippocampal gyrus
		–56	–6	–18	7.20	L	Middle temporal gyrus
		54	–2	–20	5.70	R	Middle temporal gyrus
		0	24	–20	5.36		mOFC
		–34	–10	–20	6.03	L	Hippocampus ^a
		32	–16	–18	4.96	R	Hippocampus ^b
–34	–8	–20	5.32	L	Amygdala ^c		
38	–2	–22	4.87	R	Amygdala ^d		
–38	–16	–22	4.65	L	Fusiform gyrus		
40	–26	–16	3.04	R	Fusiform gyrus		
–50	–20	–22	5.08	L	Inferior temporal gyrus		

The threshold size of activation was $p < 0.05$, corrected for multiple comparisons over the whole brain, when the height threshold was set at $t(37) > 2.72$. x, y, and z are stereotaxic coordinates (mm). R, right hemisphere; L, left hemisphere. dmPFC, dorsomedial prefrontal cortex; vmPFC, ventromedial prefrontal cortex; mOFC, medial orbitofrontal cortex; TPJ, temporoparietal junction. ^{a–d}Probability values on cytoarchitectonic maps (Amunts et al., 2005): ^a50% for the hippocampus and 40% for the amygdala; ^b90% for the hippocampus; ^c40% for the amygdala and 30% for the hippocampus; ^d50% for the amygdala.

3.2.1.3. Interactions between Tears and Sad facial expressions. The contrast of the supra-additive effect [(ST – SC) – (NT – NC)] revealed bilateral activation in vmPFC, posterior cingulate gyrus, precuneus, cuneus, parahippocampal gyrus, lingual gyrus, cerebellum, amygdala, and hippocampus. Moreover, the same contrast revealed significant activation in the left dmPFC and right mOFC (Fig. 5 and Table 4). The opposite contrast [(NT – NC) – (ST – SC)] revealed no significant activation.

Collectively, activity in the mPFC and PCC showed the main effect of tears, the main effect of sad expression, and the supra-additive effect. By contrast, activity in the TPJ showed only the main effect of tears. In order to further characterize response patterns in these regions, we conducted the following ROI analysis.

3.2.2. ROI analysis

In the whole-brain analysis, only the main effect of tears showed activation in all nodes of the core mentalizing network (the mPFC, PCC, and TPJ). Thus, we functionally defined ROIs based on this effect. In order to avoid the double-dipping problem (Kriegeskorte et al., 2009), we conducted a separate fMRI experiment to localize the mPFC, PCC, and TPJ. We found six clusters of significant activation (Supplementary Fig. 1 and Table 1). Among these, the

following five clusters corresponded to the core mentalizing network: one cluster in the superior mPFC (dmPFC and vmPFC); one cluster in the inferior mPFC (mOFC and vmPFC); one cluster in the PCC; one cluster in the left TPJ; and one cluster in the right TPJ. We used these five ROIs in the analyses (Fig. 6A).

Fig. 6B shows the contrast estimates (i.e., the activity relative to the Neutral NoTears [N] condition) for the four conditions of interest. We confirmed that the contrast estimates in the ST, SC, and NT conditions were significantly greater than the N condition in all ROIs except for the SC condition in the superior mPFC (p values < 0.05 , one-tailed one-sample t tests). More specifically, the same statistical test showed a tendency toward significance in the SC condition of the superior mPFC ($t(37) = 1.6$, $p = 0.06$).

The PCC and two clusters in the mPFC showed greater activity in the ST condition compared to the other three conditions. Two-way ANOVAs (2 levels of Tears \times 2 levels of Expressions) on the contrast estimates of these regions showed significant main effects of Tears [$F(1, 37) = 8.9$, $p < 0.01$ for the PCC; $F(1, 37) = 8.7$, $p < 0.01$ for the inferior mPFC; and $F(1, 37) = 10.3$, $p < 0.01$ for the superior mPFC] and Expressions [$F(1, 37) = 13.9$, $p < 0.01$ for the PCC; $F(1, 37) = 23.9$, $p < 0.001$ for the inferior mPFC; and $F(1, 37) = 8.6$, $p < 0.01$ for the superior mPFC]. The same ANOVA also revealed significant

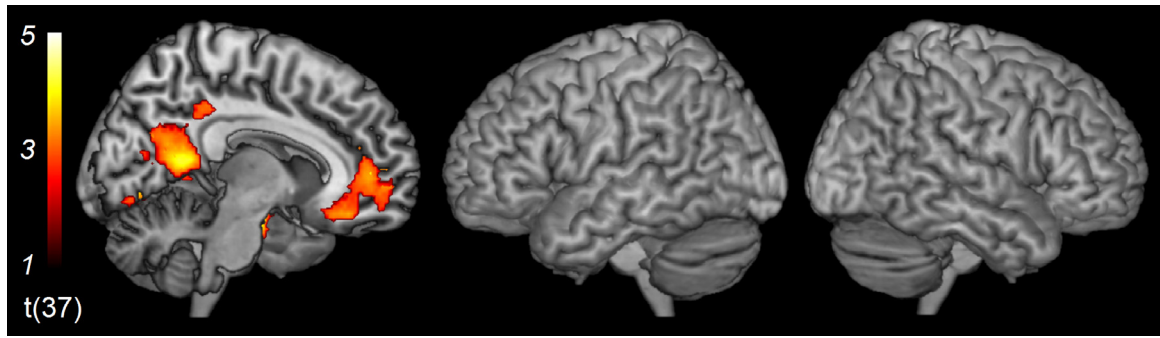


Fig. 5. The supra-additive effect between Tears and Sad expressions. The brain activation revealed by evaluating the interaction $[(ST - SC) - (NT - NC)]$ is shown. The activation was thresholded at $p < 0.05$, corrected for multiple comparisons over the whole brain, with the height threshold set at $t(37) > 2.72$ (corresponding to uncorrected $p < 0.005$).

Table 4

The supra-additive effect between Tears and Sad expressions.

Spatial extent test		MNI coordinates (mm)			t-value(37)	Location			
Cluster size (mm ³)	p-Values	x	y	z		Side	Area		
Interaction [(ST – SC) – (NT – NC)] (Supra-additive effect, Fig. 5)									
17,864	<0.001	–8	–46	30	3.72	L	Posterior cingulate gyrus		
		8	–42	34	3.05	R	Posterior cingulate gyrus		
		–10	–56	22	4.01	L	Precuneus		
		12	–50	10	4.51	R	Precuneus		
		0	–68	20	3.56		Cuneus		
		16	–40	2	4.09	R	Parahippocampal gyrus		
		0	–74	0	5.00		Lingual gyrus		
		0	–72	–10	2.75		Cerebellum		
		20	–10	–12	3.12	R	Amygdala/Hippocampus ^a		
		22	–14	–12	4.14	R	Hippocampus ^b		
		11,440	<0.001	–4	44	22	3.21	L	dmPFC
				–10	44	4	4.45	L	vmPFC
				12	46	4	3.81	R	vmPFC
				10	36	–16	3.94	R	mOFC
4088	<0.05	–34	–30	–14	4.24	L	Parahippocampal gyrus		
		–30	–10	–16	6.23	L	Amygdala ^c		
		–30	–12	–16	7.16	L	Hippocampus ^d		
Interaction [(NT – NC) – (ST – SC)]					n.s.				

The threshold size of activation was $p < 0.05$, corrected for multiple comparisons over the whole brain, when the height threshold was set at $t(37) > 2.72$. x, y, and z are stereotaxic coordinates (mm). R, right hemisphere; L, left hemisphere. dmPFC, dorsomedial prefrontal cortex; vmPFC, ventromedial prefrontal cortex; mOFC, medial orbitofrontal cortex. ^{a-d}Probability values on cytoarchitectonic maps (Amunts et al., 2005): ^a40% for the amygdala and 40% for the hippocampus; ^b50% for the hippocampus; ^c40% for the amygdala and 30% for the hippocampus; ^d60% for the hippocampus and 30% for the amygdala. n.s. indicates that no significant activation was found.

interactions between the two factors in the PCC [$F(1, 37) = 10.4$, $p < 0.01$] and inferior mPFC [$F(1, 37) = 6.2$, $p < 0.05$], and an interaction showing a trend toward significance in the superior mPFC [$F(1, 37) = 3.7$, $p = 0.06$]. Post hoc pairwise comparisons (with a Bonferroni correction) in these regions showed significantly greater contrast estimates in the ST than the SC condition (p values < 0.01), whereas there were no such differences between the NT and NC conditions (p values > 0.2).

In contrast to the midline regions, we observed no such differences in the TPJ. The same two-way ANOVAs (2 levels of Tears \times 2 levels of Expressions) revealed neither significant main effects (p values > 0.06) nor interactions (p values > 0.3). We conducted a supplementary analysis to examine further whether the spatial patterns of activation in the TPJ were identical between tears and circles (see Supplementary Materials and Methods). We found that spatial patterns were more similar within the tear condition than between the tear and circle conditions, suggesting that activation patterns in the TPJ are not identical between the two (Supplementary Fig. 2).

4. Discussion

In the present study, the mPFC and PCC showed a supra-additive effect between sad facial expressions and the presence of tears, whereas the TPJ showed no such effect.

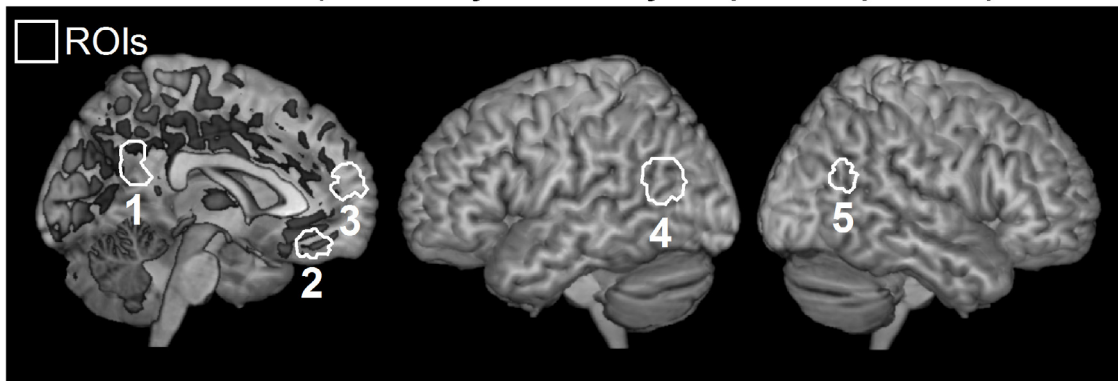
4.1. Behavioral performance

We confirmed the tears effect (Provine et al., 2009) by showing that the observation of tears increased ratings of sadness compared with control conditions in which there were no tears (no tears and circles conditions; Fig. 2). We also found interaction effects between sad facial expressions and tears: the tears effect was smaller for sad expressions than for neutral expressions. In other words, a signal of sadness shows a reduced effect when another indicator of sadness is already present. This sub-additive effect might be explained by Weber's law, which states that the change in stimulus intensity that can be discriminated is a constant fraction of the intensity of the original stimulus (e.g., facial expressions, Gao et al., 2013).

4.2. Tears effect in the mPFC, PCC, and TPJ

We found that all nodes of the core mentalizing network (mPFC, PCC, and TPJ) were activated when rating the sadness of faces with tears compared to faces without tears (Figs. 3 and 6B). To the best of our knowledge, only one study has examined the brain activity of the tears effect. More specifically, Hendriks et al. (2007) examined an early event-related potential (ERP) component (N170) in response to faces when the subject observed crying (with tears)

A: ROI definition (functionally localized by a separate experiment)



1. PCC [Center of ROI: -12 -54 26] 2. Inferior mPFC [4 42 -20] 3. Superior mPFC [-10 60 14]
4. Left TPJ [-40 -64 18] 5. Right TPJ [48 -58 26]

B: ROI analysis

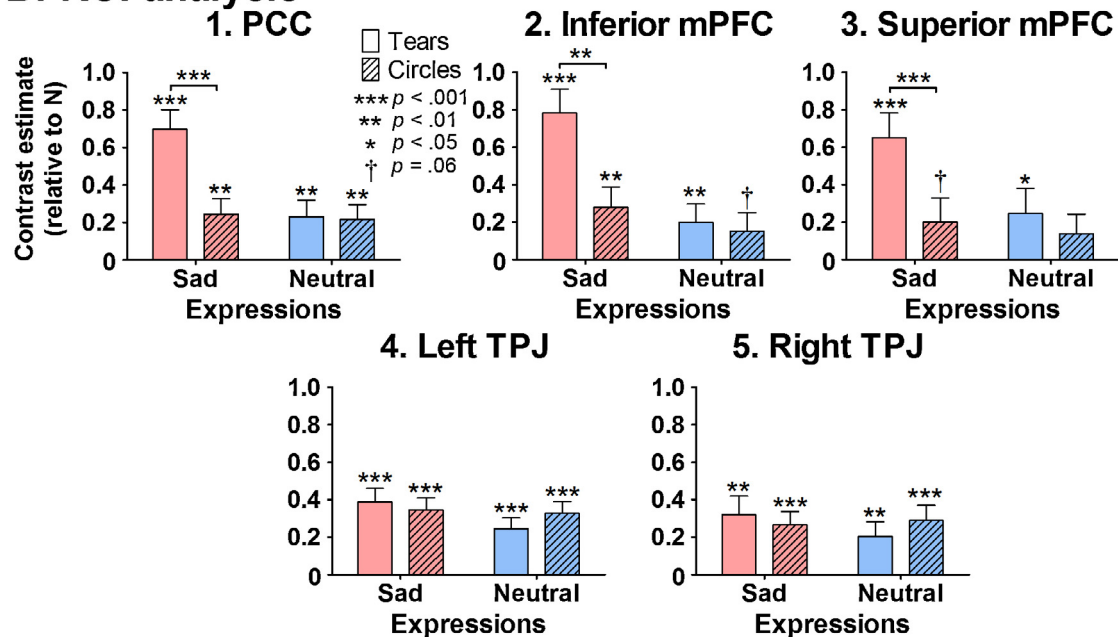


Fig. 6. ROI analysis. (A) ROI definition. ROIs are indicated by the white lines. These regions were defined by comparing Tears minus NoTears [(ST + SC) – (NT + N)] in a separate experiment (see Supplementary Fig. 1 and Table 1 for more information). (B) ROI analysis. We averaged the contrast estimates of all voxels in each ROI and examined their patterns across the four conditions of interest (ST, SC, NT, and NC). Data are presented as the mean \pm SEM. Asterisks above each bar indicate the results of one-tailed one-sample *t* tests on the contrast estimate (relative to the N condition). Asterisks between bars indicate the statistical significance of the post hoc pairwise comparisons (with the Bonferroni correction). Two-way ANOVAs on these contrast estimates showed significant interactions (the supra-additive effect) in the PCC and inferior mPFC (*p* values < 0.05). The interaction term in the superior mPFC showed a trend toward significance (*p* = 0.06).

and other facial expressions. However, neither the latency nor the amplitude of the ERPs differed between crying and other facial expressions. The present study has identified the neural substrates underlying the tears effect.

4.3. Supra-additive effects between tears and facial expressions in the mPFC and PCC

We found that the inferior mPFC (covering the mOFC and vmPFC) and the PCC showed not only main effects of facial expression and tears, but also supra-additive effects between them.

As in the field of multisensory research, our result indicates that the mPFC and PCC are involved in integrating tears and facial expressions for the purpose of inferring others' sadness. In other words, these regions might be engaged in combining the perceived social signals to infer the most likely extent of others' sadness.

Previous lesion studies showed that damage to the orbitofrontal cortex (OFC), a part of the mPFC, produces impairments in the recognition of social signals involving emotional facial expressions (Hornak et al., 1996; Blair and Cipolotti, 2000; Rolls, 2004; Dal Monte et al., 2013; Willis et al., 2014) and emotional vocal expressions (Hornak et al., 1996, 2003; Rolls, 2004). In accord with these findings, previous neuroimaging studies have indicated that the mPFC contains abstract representations of others' emotional states regardless of the type of social signal (Peelen et al., 2010; Skerry and Saxe, 2014). However, as each type of social signal was presented separately in these studies, it was unclear whether the mPFC showed an interaction effect between multiple social signals. Moreover, unlike the mPFC, the role of the PCC has been poorly investigated in the context of mentalizing. The current study revealed that the mPFC and PCC showed an interaction effect between distinct social signals during affective mentalizing,

providing more direct evidence for the integration of social signals in these regions.

One explanation of our result is that the integration process is conducted in other brain regions, and the supra-additive effect in the mPFC and PCC represents the extent of others' sadness provided by such an integration process. However, this interpretation was not supported by two findings: first, only the mPFC and PCC consistently met the criteria of convergence (i.e., were activated by each social signal) and interaction; and second, the sadness rating showed the sub-additive effect, but not the supra-additive effect (Fig. 2). Thus, it is unlikely that the supra-additive effect in these regions simply represents the extent of sadness. Rather, it is better explained by the hypothesis that integrating two different social signals (for the purpose of affective mentalizing) imposes greater processing demands in the mPFC and PCC than individual signals.

The core mentalizing network (such as the mPFC and PCC) and the human homologue of the mirror-neuron system (e.g., the inferior frontal gyrus and inferior parietal lobule) are both active during the recognition of others' facial emotions (Phan et al., 2002; Carr et al., 2003; Winston et al., 2003; Lennox et al., 2004; Vytal and Hamann, 2010; Kitada et al., 2013). As compared to the mirror-neuron system, the core mentalizing network seems to be active when observers reflect on the cause of the behavior – e.g., why is this person shedding tears? (Van Overwalle and Baetens, 2009). The core mentalizing network is also proposed to be a part of the social “reflective system” (C system), a slow system that is responsible for taking situational constraint information and other prior knowledge into account for mentalizing (Satpute and Lieberman, 2006). In order to determine that the person shedding tears is sad, we rely on knowledge based on previous experience. Consistent with this view, the PCC is associated with long-term memory (Minoshima et al., 1997; Ranganath et al., 2004; Wagner et al., 2005; Cavanna and Trimble, 2006; Matsuda, 2007). For instance, Alzheimer's disease (AD) is characterized not only by medial temporal lobe (MTL) atrophy, but also by a reduction of glucose metabolism in the cingulo-parietal cortex, including the precuneus (Matsuda, 2007). Activity in the precuneus is reduced in patients with very-early-stage AD who exhibit only memory impairment, without general cognitive decline (Minoshima et al., 1997). The default mode network (DMN), including the PCC and mPFC, is often associated with mind wandering, which can result in the retrieval of an episodic memory (Mason et al., 2007; Spreng et al., 2009). In the present study, the PCC showed activation not only in the presence of tears, but also during the observation of sad facial expressions (relative to neutral expressions; Figs. 4 and 6B). Therefore, the PCC might be involved in the retrieval of the social meaning of tears (i.e., as a symbol of sadness) from long-term memory, and the integration of the retrieved social knowledge with the sad facial expression.

In addition to the PCC, we found the supra-additive effect in the inferior mPFC, including the mOFC and vmPFC. Subregions in the mPFC are thought to play distinct but complementary roles in mentalizing (Amodio and Frith, 2006; Krueger et al., 2009). Krueger et al. (2009) proposed that the inferior mPFC supports inferences about the likely affective response and reward value accompanying goal achievement. According to this hypothesis, the process of integrating tears and facial expressions in this region might reflect the evaluation of the state of sadness (i.e., how sad is this person?). Lesions in the OFC can lead to abnormal social judgments in response to emotional faces (Willis et al., 2010). More specifically, the subjects in this study were presented with faces portraying emotional expressions and asked to imagine whether they would approach them to ask for directions. Compared to control (intact) subjects and patients with damage to frontal regions sparing the OFC, the patients with damage to the OFC tended to have abnormal approachability judgments: OFC patients rated faces displaying negative emotional expressions as significantly more approachable

than the other subject groups. We also found that the superior mPFC showed a tendency toward the supra-additive effect (Fig. 6B). The superior mPFC is thought to support inferences about the likely actions performed by others for goal achievement (i.e., why is this person shedding tears?) (Krueger et al., 2009). As such, it is possible that, along with the PCC, these subregions in the mPFC might work in concert to infer another's state of sadness at an abstract level.

We also observed the supra-additive effect in areas of the limbic system such as the MTL and amygdala (Table 4). The supra-additive effect in the MTL is consistent with our speculation that the social knowledge of tears is retrieved from long-term memory and integrated with sad facial expressions (Eldridge et al., 2000; Miyashita, 2004). The amygdala is considered as a part of the social “reflexive system” (X system), which automatically and quickly evaluates others' behavior (Satpute and Lieberman, 2006). Thus, this reflexive system (X system) might be engaged in the integration of social signals directly (Morris et al., 1998) or indirectly via the top-down modulation from parts of the reflective system (C system), such as the mPFC and PCC (Ochsner et al., 2002; Pessoa et al., 2002). However, unlike the mPFC and PCC, the MTL and amygdala were not consistently activated by the presence of tears in the two experiments; they were not activated in the separate experiment (Supplementary Fig. 1 and Table 1). Therefore, further studies are necessary to examine whether the MTL and amygdala are important for the integration of tears and facial expressions.

4.4. No interaction effect in the TPJ

Like the mPFC and PCC, the TPJ showed greater activation when viewing faces with tears than without tears, which was a more liberal control than circles (Figs. 3 and 6A). However, we observed no interaction between facial expressions and tears (Fig. 6B). These results indicate that the TPJ is engaged not in the integration of signals related to tears and facial expressions, but rather in the processing of tears.

The TPJ is thought to be a hub of diverse functions, including perceptual/motor reorienting and theory of mind (Cabeza et al., 2012). It has been proposed that one of the general functions of the TPJ is to detect a mismatch between our expectations and actual outcomes (Corbetta et al., 2008; Koster-Hale and Saxe, 2013). More specifically, the TPJ is activated when a target is presented in an unexpected location in spatial attention tasks (Posner paradigm; Corbetta and Shulman, 2002). False-belief stories, which are often used in mentalizing tasks, require processing information detected outside the main focus of attention (Cabeza et al., 2012). Shedding tears (in adults) and circles are rarely observed in our daily life, whereas the activation pattern between tears and circles differed in the TPJ. Therefore, it is possible that the TPJ is involved in detecting and perceiving unusual objects such as tears.

4.5. Limitations

Two limitations must be considered. First, we utilized pictures of different individuals between the sad and neutral conditions. It is unlikely that the supra-additive effect is also affected by the difference in facial identity, because this factor is subtracted out in the supra-additive effect $[(ST - SC) - (NT - NC)]$. However, we cannot rule out the possibility that greater activity in sad expressions (relative to neutral expressions) can be partially explained by different facial identity. Second, we used only female subjects, because they tend to react to crying people with more sympathy and support than males (Cretser et al., 1982). However, future studies should test whether this finding can be generalized to male subjects, and examine the integration of facial expressions and tears in genders that are different from the subjects (the cross-gender effect).

5. Conclusions

The present study investigated which nodes of the core mentalizing network are involved in the integration of tears and facial expressions that are used to infer the extent of another's sadness. We found that the mPFC and PCC showed a supra-additive effect between tears and facial expressions. In contrast, the TPJ showed no such effect. These results indicate that the mPFC and PCC are involved in integrating distinct social signals to represent others' sadness at an abstract level. These results highlight the differences in the contributions of the mPFC, PCC, and TPJ to affective mentalizing.

Acknowledgments

The authors declare no competing financial interests. This work was supported by the Sasakawa Scientific Research Grant from The Japan Science Society to HKT; by a Grant-in-Aid for Young Scientists (B) (number 25871059) from the Japan Society for the Promotion of Science, and a Grant-in-Aid for Scientific Research on Innovative Areas, "Face perception and recognition" (number 23119727) from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan to RK; and by a Grant-in-Aid for Scientific Research S (number 21220005) A(number 15H0186) and "Development of biomarker candidates for social behavior" carried out under the Strategic Research Program for Brain Sciences from the MEXT of Japan to NS.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.neures.2015.07.005>

References

- Amodio, D.M., Frith, C.D., 2006. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277.
- Amunts, K., Kedo, O., Kindler, M., Pieperhoff, P., Mohlberg, H., Shah, N.J., Habel, U., Schneider, F., Zilles, K., 2005. Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. *Anat. Embryol. (Berl.)* 210, 343–352.
- Atique, B., Erb, M., Gharabaghi, A., Grodd, W., Anders, S., 2011. Task-specific activity and connectivity within the mentalizing network during emotion and intention mentalizing. *Neuroimage* 55, 1899–1911.
- Balsters, M.J., Krahmer, E.J., Swerts, M.G., Vingerhoets, A.J., 2013. Emotional tears facilitate the recognition of sadness and the perceived need for social support. *Evol. Psychol.* 11, 148–158.
- Barbey, A.K., Krueger, F., Grafman, J., 2009. Structured event complexes in the medial-prefrontal cortex support counterfactual representations for future planning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1291–1300.
- Blair, R.J., Cipolletti, L., 2000. Impaired social response reversal: a case of 'acquired sociopathy'. *Brain* 123, 1122–1141.
- Cabeza, R., Ciaramelli, E., Moscovitch, M., 2012. Cognitive contributions of the ventral parietal cortex: an integrative theoretical account. *Trends Cogn. Sci.* 16, 338–352.
- Calvert, G.A., Campbell, R., Brammer, M.J., 2000. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr. Biol.* 10, 649–657.
- Carr, L., Iacoboni, M., Dubeau, M.C., Mazziotta, J.C., Lenzi, G.L., 2003. Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5497–5502.
- Cavanna, A.E., Trimble, M.R., 2006. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129, 564–583.
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.
- Corbetta, M., Patel, G., Shulman, G.L., 2008. The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324.
- Corradi-Dell'Acqua, C., Hofstetter, C., Vuilleumier, P., 2014. Cognitive and affective theory of mind share the same local patterns of activity in posterior temporal but not medial prefrontal cortex. *Soc. Cogn. Affect. Neurosci.* 9, 1175–1184.
- Cretser, G.A., Lombardo, W.K., Lombardo, B., Mathis, S., 1982. Reactions to men and women who cry: a study of sex differences in perceived societal attitudes versus personal attitudes. *Percept. Mot. Skills* 55, 479–486.
- Dal Monte, O., Krueger, F., Solomon, J.M., Schintu, S., Knutson, K.M., Strenziok, M., Pardini, M., Leopold, A., Raymond, V., Grafman, J., 2013. A voxel-based lesion study on facial emotion recognition after penetrating brain injury. *Soc. Cogn. Affect. Neurosci.* 8, 632–639.
- Decety, J., Lamm, C., 2007. The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist* 13, 580–593.
- Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335.
- Eldridge, L.L., Knowlton, B.J., Furmanski, C.S., Bookheimer, S.Y., Engel, S.A., 2000. Remembering episodes: a selective role for the hippocampus during retrieval. *Nat. Neurosci.* 3, 49–52.
- Evans, A.C., Kamber, M., Collins, D.L., MacDonald, D., 1994. An MRI-based probabilistic atlas of neuroanatomy. In: Shorvon, S.D., Fish, D.R., Andermann, F., Bydder, G.M., Stefan, H. (Eds.), *Magnetic Resonance Scanning and Epilepsy*, vol. 264. Springer, New York, pp. 263–274.
- Friston, K.J., Jezzard, P., Turner, R., 1994a. Analysis of functional MRI time-series. *Hum. Brain Mapp.* 1, 153–171.
- Friston, K.J., Worsley, K.J., Frackowiak, R.S., Mazziotta, J.C., Evans, A.C., 1994b. Assessing the significance of focal activations using their spatial extent. *Hum. Brain Mapp.* 1, 210–220.
- Friston, K.J., Ashburner, J., Frith, C.D., Heather, J.D., Frackowiak, R.S.J., 1995. Spatial registration and normalization of images. *Hum. Brain Mapp.* 2, 165–188.
- Friston, K.J., Holmes, A., Poline, J.B., Price, C.J., Frith, C.D., 1996. Detecting activations in PET and fMRI: levels of inference and power. *Neuroimage* 4, 223–235.
- Friston, K.J., Glaser, D.E., Henson, R.N., Kiebel, S., Phillips, C., Ashburner, J., 2002. Classical and Bayesian inference in neuroimaging: applications. *Neuroimage* 16, 484–512.
- Friston, K.J., Ashburner, J., Kiebel, S.J., Nichols, T.E., Penny, W.D., 2007. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, London.
- Frith, U., Frith, C.D., 2003. Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 459–473.
- Gao, X., Maurer, D., Nishimura, M., 2013. Altered representation of facial expressions after early visual deprivation. *Front. Psychol.* 4, 878.
- Goel, V., Grafman, J., Sadato, N., Hallett, M., 1995. Modeling other minds. *Neuroreport* 6, 1741–1746.
- Hendriks, M.C., van Boxtel, G.J., Vingerhoets, A.J., 2007. An event-related potential study on the early processing of crying faces. *Neuroreport* 18, 631–634.
- Holmes, A.P., Friston, K.J., 1998. Generalizability, random effects and population inference. *Neuroimage* 7, 5754.
- Hornak, J., Rolls, E.T., Wade, D., 1996. Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. *Neuropsychologia* 34, 247–261.
- Hornak, J., Bramham, J., Rolls, E.T., Morris, R.G., O'Doherty, J., Bullock, P.R., Polkey, C.E., 2003. Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain* 126, 1691–1712.
- Kitada, R., Okamoto, Y., Sasaki, A.T., Kochiyama, T., Miyahara, M., Lederman, S.J., Sadato, N., 2013. Early visual experience and the recognition of basic facial expressions: involvement of the middle temporal and inferior frontal gyri during haptic identification by the early blind. *Front. Hum. Neurosci.* 7, 7.
- Koster-Hale, J., Saxe, R., 2013. Theory of mind: a neural prediction problem. *Neuron* 79, 836–848.
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540.
- Krueger, F., Barbey, A.K., Grafman, J., 2009. The medial prefrontal cortex mediates social event knowledge. *Trends Cogn. Sci.* 13, 103–109.
- Lennox, B.R., Jacob, R., Calder, A.J., Lupson, V., Bullmore, E.T., 2004. Behavioural and neurocognitive responses to sad facial affect are attenuated in patients with mania. *Psychol. Med.* 34, 795–802.
- Mason, M.F., Norton, M.I., Van Horn, J.D., Wegner, D.M., Grafton, S.T., Macrae, C.N., 2007. Wandering minds: the default network and stimulus-independent thought. *Science* 315, 393–395.
- Matsuda, H., 2007. Role of neuroimaging in Alzheimer's disease, with emphasis on brain perfusion SPECT. *J. Nucl. Med.* 48, 1289–1300.
- Meredith, M.A., Stein, B.E., 1983. Interactions among converging sensory inputs in the superior colliculus. *Science* 221, 389–391.
- Minoshima, S., Giordani, B., Berent, S., Frey, K.A., Foster, N.L., Kuhl, D.E., 1997. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann. Neurol.* 42, 85–94.
- Mitchell, J.P., 2008. Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cereb. Cortex* 18, 262–271.
- Miyashita, Y., 2004. Cognitive memory: cellular and network machineries and their top-down control. *Science* 306, 435–440.
- Morris, J.S., Ohman, A., Dolan, R.J., 1998. Conscious and unconscious emotional learning in the human amygdala. *Nature* 393, 467–470.
- Murube, J., Murube, L., Murube, A., 1999. Origin and types of emotional tearing. *Eur. J. Ophthalmol.* 9, 77–84.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Ochsner, K.N., Bunge, S.A., Gross, J.J., Gabrieli, J.D., 2002. Rethinking feelings: an fMRI study of the cognitive regulation of emotion. *J. Cogn. Neurosci.* 14, 1215–1229.
- Peelen, M.V., Atkinson, A.P., Vuilleumier, P., 2010. Supramodal representations of perceived emotions in the human brain. *J. Neurosci.* 30, 10127–10134.

- Perry, A., Shamay-Tsoory, S., 2013. *Understanding Emotional and Cognitive Empathy: A neuropsychological perspective*. In: Baron-Cohen, S., Lombardo, M., Tager-Flusberg, H. (Eds.), *Understanding Other Minds: Perspectives from Developmental Social Neuroscience*. Oxford University Press, London, pp. 178–194.
- Pessoa, L., McKenna, M., Gutierrez, E., Ungerleider, L.G., 2002. Neural processing of emotional faces requires attention. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11458–11463.
- Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I., 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage* 16, 331–348.
- Provine, R.R., Krosnowski, K.A., Brocato, N.W., 2009. Tearing: breakthrough in human emotional signaling. *Evol. Psychol.* 7, 52–56.
- Raij, T., Uutela, K., Hari, R., 2000. Audiovisual integration of letters in the human brain. *Neuron* 28, 617–625.
- Rolls, E.T., 2004. The functions of the orbitofrontal cortex. *Brain Cogn.* 55, 11–29.
- Ranganath, C., Cohen, M.X., Dam, C., D'Esposito, M., 2004. Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *J. Neurosci.* 24, 3917–3925.
- Samson, D., Apperly, I.A., Chiavarino, C., Humphreys, G.W., 2004. Left temporoparietal junction is necessary for representing someone else's belief. *Nat. Neurosci.* 7, 499–500.
- Saxe, R., Powell, L.J., 2006. It's the thought that counts: specific brain regions for one component of theory of mind. *Psychol. Sci.* 17, 692–699.
- Satpute, A.B., Lieberman, M.D., 2006. Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain Res.* 1079, 86–97.
- Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E.N., Saxe, R., 2009. Distinct regions of right temporo-parietal junction are selective for theory of mind and exogenous attention. *PLoS ONE* 4, e4869.
- Shamay-Tsoory, S.G., Aharon-Peretz, J., Perry, D., 2009. Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain* 132, 617–627.
- Shattuck, D.W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K.L., Poldrack, R.A., Bilder, R.M., Toga, A.W., 2008. Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage* 39, 1064–1080.
- Skerry, A.E., Saxe, R., 2014. A common neural code for perceived and inferred emotion. *J. Neurosci.* 34, 15997–16008.
- Spreng, R.N., Mar, R.A., Kim, A.S., 2009. The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *J. Cogn. Neurosci.* 21, 489–510.
- Stevenson, R.A., Kim, S., James, T.W., 2009. An additive-factors design to disambiguate neuronal and areal convergence: measuring multisensory interactions between audio, visual, and haptic sensory streams using fMRI. *Exp. Brain Res.* 198, 183–194.
- Van Overwalle, F., 2009. Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* 30, 829–858.
- Van Overwalle, F., Baetens, K., 2009. Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage* 48, 564–584.
- Vytal, K., Hamann, S., 2010. Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis. *J. Cogn. Neurosci.* 22, 2864–2885.
- Wagner, A.D., Shannon, B.J., Kahn, I., Buckner, R.L., 2005. Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* 9, 445–453.
- Willis, M.L., Palermo, R., Burke, D., McGrillen, K., Miller, L., 2010. Orbitofrontal cortex lesions result in abnormal social judgments to emotional faces. *Neuropsychologia* 48, 2182–2187.
- Willis, M.L., Palermo, R., McGrillen, K., Miller, L., 2014. The nature of facial expression recognition deficits following orbitofrontal cortex damage. *Neuropsychology* 28, 613–623.
- Winston, J.S., O'Doherty, J., Dolan, R.J., 2003. Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *Neuroimage* 20, 84–97.
- Worsley, K.J., Friston, K.J., 1995. Analysis of fMRI time-series revisited-again. *Neuroimage* 2, 173–181.